PERFORMANCE INDICATORS FOR URBAN PUBLIC TRANSPORT SYSTEMS WITH A FOCUS ON TRANSPORT POLICY EFFECTIVENESS ISSUES

Marco Diana - Politecnico di Torino, Dipartimento di Idraulica, Trasporti e Infrastrutture Civili - Corso Duca degli Abruzzi, 24 - 10129 Torino - Italy - Ph. +39 011 564 5638 - Fax +39 011 564 5699 - marco.diana@polito.it

Cinzia Daraio - University of Bologna, CIEG - Centro Studi di Ingegneria Economico-Gestionale, Department of Management - Via Umberto Terracini, 28 - 40131 Bologna - Italy - Ph. +39 051 2090210 - Fax +39 051 2090222 - cinzia.daraio@unibo.it

ABSTRACT

The present paper reviews the methodologies that have been proposed to assess the performances of urban public transport systems, in order to identify possible research gaps. An evaluation framework is proposed to adequately address the specificities of public transport as a field of intervention of public powers, that make it different from other subsidized sectors such as health care or education. According to such framework, it is found that efficiency issues are the most studied ones, dealing with various aspects of service production with a managerial perspective, whereas customer-related quality issues have come into play in more recent years. However, the paper argues that the point of view of policy makers does not necessarily coincide with any of these two. Related transport policy effectiveness issues have been less consistently studied and probably need to be pushed up in the research agenda. The paper ends by proposing a preliminary set of new transport effectiveness indicators to properly evaluate the contribution of transit to improve a transport system on three of the aspects that are usually at the core of any policy action, namely accessibility improvement, modal diversion and environmental impacts.

Keywords: Public transport, performance evaluation, effectiveness indicators

INTRODUCTION

Keep public transport up and running is one of the most significant sources of expenditure for any territorial public body (central state, regions, provinces and municipalities) in most countries. The reasons that implicitly or explicitly are at the basis of such a huge effort usually go much beyond transport-related matters. Beyond improving the technical efficiency and the quality of the offer of the public transport system itself, stakeholders are in fact usually interested in positively influencing some dynamics in a wider range of ambits, whenever they invest in such services. According to the specific context, expected benefits might go from reducing pollution to improving social equality. Most remarkably, these “derived effects” often constitute the primary reason for justifying such interventions, both in current political discourses but also at the more technical level of the transport planning documents. In such cases, the focus is not so much on the service that is provided to a set of potential customers, but on the global impacts through the modification of environmental footprints, land use patterns or territorial accessibility.

The present paper deals with evaluation issues of public transport by explicitly taking into consideration such distinguishing feature of transit services. In this sense, we note that public transport is obviously different from those unsubsidized economic activities that are only subject to market laws, but also from other sectors of public intervention such as health care or education, where the primary goal is more straightforward and only the direct benefits for their users have a substantial importance. Our main point is therefore that the importance of the above mentioned “derived effects” of public transport systems should be reflected in their evaluation methods. This has been only partially acknowledged in past research efforts.

The public transport evaluation perspective that we consider here is related to day-to-day operations rather than to planning, capital investments and other strategic decisions. We believe in fact that evaluation practices are better established in the latter case, where a wealth of methods are available and routinely used. On the contrary, the evaluation of the daily service operations and the assessment of the effectiveness of related subsidies is a continuous process that requires an organisation effort and dedicated resources inside the firm. Therefore, it is essential to raise the awareness on the importance of this issue among the interested parties to effectively monitor service operations. Too often such evaluations tend in fact to become a formal exercise, that is even neglected in many countries.

Evaluation activities of the service operation can be carried out through several different methods. In the present research we limit ourselves to examine the most commonly used one at the practitioner level, i.e. the use of a set of performance indicators. The first objective of the present paper is therefore to deeply review the existing literature related to public transport performance indicators, and to assess to what extent they can be used for a transport policy effectiveness analysis. While “effectiveness” indicators are widely available in the literature, the definitions of effectiveness that are adopted in these works are in fact often different from the one we need to consider here. In order to fully meet our research goal and truly embrace the point of view of the decision maker, we build a more general framework, in which several different kinds of indicators are needed in order to cover the above mentioned socially relevant aspects, and to frame them within a unique evaluation activity. The second objective of the work will therefore be to propose some indicators that are relevant to such evaluation perspective.

Although most of the contents of this paper can be referred to any kind of public transport, in the following we mainly refer to urban public transport systems, to which are directed most public resources. The following section of the paper is devoted to set up our evaluation framework and to discuss the different meanings of the concept of effectiveness that have been previously developed, in order to provide a definition of effectiveness that is consistent with our research perspective. Then we review the relevant literature in this field, where the indicators that have been proposed by different researchers over the decades will be classified according to the perspective we defined. The following section will present some indicators among the reviewed ones that are most useful in a transport policy effectiveness study, and the final section will offer some concluding remarks and give subsequent research perspectives.

TRANSPORT POLICY EFFECTIVENESS EVALUATION FRAMEWORK AND DEFINITIONS

Different approaches are used both at the research and at the practitioner level to evaluate public transport. These divergences are often more due to the disciplinary background of the analysts, than to the deliberate choice of using the most appropriate working method under a given set of circumstances. At least two different streams of research can be detected concerning this point. On one hand, civil engineers and transport planners are usually more interested in studying the technical performances of the system itself, while being conscious of the different and far-reaching implications that often underlie any investment in a public transport system. This generally leads to the definition of a set of performance indicators, since this very flexible method allows for jointly considering heterogeneous kinds of data (for example, public subsidies, commercial speed and decrease of pollutants emissions) in a rather straightforward way on an analytical point of view, often involving simple mathematical operations. Analysts from those research areas should also be aware of the fact that the particular contexts in which the different transport systems operate, particularly concerning population size, socioeconomic characteristics and land use patterns, make it difficult to carry out a comparative assessment of the performances of different systems. Early studies dealing with this problem include the works of Mundle and Cherwony (1980), Giuliano (1981), Vaziri and Deacon (1984) and Fielding et al. (1985b).
On the other hand, quantitative economic analysts tend to apply efficiency analyses as they do in other disciplinary fields. Their methodologies are very insightful, allowing one to evaluate how well the system is operating, if the resources could be better used etc. As such, these tools are very useful from a public transport operator viewpoint, like for any other firm, but from the more general point of view of the stakeholder they should be integrated according to the above described perspective. Also, particular care should be taken in comparative performance assessment exercises, in order to consider only those systems that are effectively comparable. Recalling the above mentioned influence of territorial factors, there is in fact not just one single market with several operators to compare, but as many markets as public transit systems on the different territories, each territory possibly hosting just one operator.
The concept of economic efficiency has been declined over the years in different ways to make analyses more comprehensive. Farrell (1957) -extending the work initiated by Koopmans and Debreu- distinguished between technical and allocative efficiency. The concept of technical efficiency is associated with the production frontier which measures the success in producing the maximum output from a given set of inputs; whereas the concept of allocative efficiency is related to the choice of an optimal set of inputs, given the inputs price. Generally, in the public sector, prices of inputs and outputs are not always available, and for

that reason, the technical efficiency concept in the most used in this framework. Reviews on the body of work concerning economic efficiency studies of transport systems can be found in De Borger et al. (2002) and Brons et al. (2005) concerning public transit, and in Oum et al. (1999) for the rail sector.

A simple look at the extensive literature in those two disciplinary fields (transport planning on one side and economics on the other) allows one to conclude that they essentially proceed along parallel routes to substantially study the same problem, with few occasion of exchanges both at the scientific and at the practical implementation levels. Yet an evaluation methodology that encompasses these two different approaches to pick up their respective strengths would improve the current state of the art, as argued by Daraio and Diana (2009). In the following, albeit taking the transport planning perspective and therefore focusing on performance indicators, we will explicitly take into consideration the possibility of such an integration, which constitutes a desired future development of our work.

We would preliminarily like to clarify some definitions of the terms that are used in the public transport evaluation literature. The most common ones are performance, efficiency and effectiveness, for which several definitions have been given (Talley and Anderson, 1981). ***Performance*** is usually a more general word that is used to encompass the different points of view that can be considered. In other words, the assessment of the performance of a system can be done by considering several different aspects, including efficiency and effectiveness. In this paper we will use the word “performance” in such general terms. On the other hand, given the above mentioned intensive research in economics, and in particular in public economics where prices of inputs and outputs are not always available, we have the more precisely defined concept of technical efficiency. ***Technical Efficiency*** is therefore considered as the ratio between the outputs (produced quantities) of a process and the required inputs (resources needed) to produce the outputs.

Effectiveness is instead a concept with many more facets and less consistent definitions across different authors, since disparate research perspectives contribute in shaping it. We present in table 1 some of the definitions of effectiveness that can be found in the public transport evaluation literature. It is easy to note differences not only in meaning, but even in the terminology being used, that point to different disciplinary backgrounds of the various authors. In particular, the first five rows give examples of definitions that seems more in relation with an economic perspective. According to these authors, effectiveness is essentially in relation with the interaction between demand and offer in the public transport market, whereas efficiency deals more with the characteristics of the offer. On the other hand, the following five rows report definitions that essentially adopt the engineering and planning perspective, thus focussing on the transport service goals and on the extent to which they are reached. However each definition differs in identifying the entity that should express these goals. Finally, the last three rows report more recent definitions where the user’s point of view seems to prevail. Therefore, we see that the notion of effectiveness has also evolved over time, beyond the differences among different authors and their respective disciplines.

Table 1 – Definitions of public transport effectiveness according to different Authors

References	***Definition***
Economic dimension	
McCrosson (1978); Talley	Concerned with consumed output

(1988)	
Fielding et al. (1978)	Comparison of produced output to intended output or objectives
Giuliano (1981)	The extent of service consumption
Keck et al. (1980); Fielding et al. (1985a); Takyi (1993); Lem et al. (1994)	The degree to which outputs are consumed (referred by Fielding et al., 1985a et Lem et al., 1994 as "service effectiveness", different from "cost effectiveness")
Tulkens and Wunsch (1994)	Adequacy of the offer to the demand
Adherence to objectives	
Gleason and Barnum (1982)	The extent to which an objective has been achieved
Dajani and Gilbert (1978)	The degree to which the transit services achieves individual and community mobility goals
Khan (1981)	Ability to meet the broader service and operational objectives resulting from a combination of community, industry and government interests
Fielding and Lyons (1981)	The extent to which service consumed corresponds to government goals and objectives
Kelley and Rutherford (1982)	How well transit is meeting the goals and objectives set by government policy
Customers viewpoint	
Yeh et al. (2000)	The degree to which the transport service meets the passenger's needs
Hensher and Prioni (2002)	Effectiveness for users: service quality
TCRP (2003a)	Cost effectiveness: the ability to meet the demand given existing resources. Service effectiveness: persons transported given existing resources

To try to sort out this set of definitions, the key point is to decide which point of view we should adopt, since it is evident that the interests of different groups related to public transport are diverging. Allen and DiCesare (1976) were among the first authors that identified the following three different "actors" that are interested in public transport operations, whose different perspectives should be properly addressed in an evaluation exercise:

1. *Public bodies* subsidize the system as a tool to reach some general transport policy objectives of interest for the whole community, ranging from the reduction of the environmental impacts of transport to the promotion of social equality and the land use planning.

2. *Public transport operators*, whether they are publicly or privately owned, should adopt the point of view of every enterprise, whose objective is profit maximisation or losses minimisation.

3. *Public transport users* are the consumers, and therefore are interested in a service with the maximum quality-to-cost ratio, where the concept of quality encompasses several different characteristics of the service that will be reviewed in the next section.

Concerning the third item in the above list, we already noted that the point of view of users is being increasingly considered in more recent years, as we better show in the next section when reviewing the different indicators that have been proposed to evaluate public transport. However, considering the traveller's perspective to define effectiveness seems rather an exception. Consistently with the marketing literature in other sectors, there is in fact a more general consensus to use the term ***quality*** to express the point of view of the service users. Since the managerial perspective of the firms is well captured by the notion of efficiency, as it is apparent by its previously given definition, it seems therefore logical to link the point of view of the funding agencies and the public bodies with the notion of effectiveness. The operational definition of effectiveness that we will use is therefore the following: *the degree to which the public transport system meets the policy objectives that were set by its public funding body*. By looking at table 1, we notice that a similar definition has been for example done by Kelley and Rutherford (1982).

Relating the notion of effectiveness of public transport exclusively with the point of view of the public body that finances it, rather than with the interests of service users, is consistent with the specificity of this sector that was mentioned in the introduction. We already pointed out in fact the differences from those services on the market, where the only economic actors are firms and consumers, but also from other sectors of public intervention, where only the direct benefits to users justify public expenditure. In such cases, it is clear that the user point of view is prevalent also to evaluate effectiveness aspects from a policy perspective. However, if the decision maker is financing public transport to obtain some derived effects beyond the benefits to users, such derived effects must be considered in the evaluation exercise. Concerning this issue, Gleason and Barnum (1982) point out the ambiguities in some commonly used effectiveness indicators, providing numerical examples in which the selection of the most effective transport system could be affected by the kind of indicator being used. This is because apparently similar indicators are in reality connected to different objectives and perspectives (for example, maximizing ridership versus minimizing the cost of a ride). To avoid such ambiguities, we think that the notion of effectiveness should exclusively be linked with the objectives of the policy makers that partly finances the system, since the system operator and the users are already concerned with efficiency and quality aspects respectively.

We finally note that our definition of effectiveness could somewhat be put in relationship with the notion of ***impact***, that can in fact be defined as the effect of a given activity on the wider environment and community. Some evaluation frameworks actually distinguish between effectiveness and impacts, for example relating the former to the effects on individuals and the latter to the effects on the larger social, economic or ecologic environment (e.g. Dajani and Gilbert, 1978). However, within our framework, we want once more to stress that such far-reaching effects are usually among the goals of public transport financing, rather than derived or collateral effects that need to be somewhat managed. The effectiveness of a public transport system cannot therefore be disjointed from its impacts. This implies that impacts should be fully embedded in the transport policy effectiveness evaluation activity, possibly through appropriate indicators, as we later show.

REVIEW OF PERFORMANCE INDICATORS FOR URBAN PUBLIC TRANSPORT

Earlier indicators related to efficiency and effectiveness issues

The performance of public transport is the object of intensive research activity at least since the late Seventies. The use of indicators is the most popular tool, given its several practical advantages, including the easiness of use and the intuitive meaning, that make them quite attractive to easily convey the results of the evaluation exercise to a wider audience. Early research efforts were thus aimed at defining a set of indicators on the basis of the data that were available, and investigating issues related to the comparability of different systems. In the U.S., federal legislation set up a reporting system for all transit agencies in order to receive funding (the so-called Section 15 database), which paved the way for a comprehensive effort in evaluating existing transport systems throughout that country.

Public transport evaluation methods dating back the "deregulation age" in the Eighties were mainly derived from the private sector, thus focusing on economic efficiency issues and primarily considering the point of view of transit operators. The most important goal in those years, at least in the U.S., was in fact to maintain the economic sustainability of the service in presence of increasing operating costs and declining patronage. Efficiency was therefore the primary concern, with effectiveness issues being mostly considered as the way to ensure that the system would effectively meet the user's requirements, according to the definitions of effectiveness that have been discussed in the preceding section.

In light of the previously developed framework, we then present in table 2 a list of indicators that have been proposed in the former literature and that primarily deal with efficiency and, to a much lesser extent, effectiveness issues. This list is obviously not exhaustive, due to both incomplete review and to the fact that we omitted for example more unusual indicators, indicators related to more specific aspects of the service (such as maintenance issues) or indicators that are clearly inferior compared to other reported ones. Beyond this, only ratio indicators are reported, thus omitting the presentation of simpler quantities describing the service (such as headways, commercial speed, load factors or line capacities) that nevertheless can be considered as efficiency or effectiveness indicators as well. In an effort to present a list as much coherent as possible and to better match the different researchers' perspectives and definitions, in some cases the indicators here presented might not exactly be those proposed by the respective authors. This is also due to the fact that some papers omit to exactly define the quantities that are used to build an indicator, for example whether "costs" include both capital and operating costs, "vehicle km" include both revenue vehicle km and trips to/from the depot, or whether "passenger places" include just seats or also standing places. In this latter case we always indicate seats in the table, since this makes no conceptual difference in the evaluation exercise.

The first column of the table reports the data that are needed to compute each indicator. The most commonly used abbreviations are explained in appendix 1, whereas some other quantities specific for a given indicator are defined in the last column of the table. The reader is referred to the papers mentioned in the last column for a more complete discussion of each measure.

Table 2 – Performance indicators for public transport

Indicator	*Measurement Unit*	*References and Notes*
1. Technical efficiency		
Exp / Veh km Op Exp / Veh km Exp / Rev veh km Op Exp / Rev veh km	€ / veh km	McCrosson (1978); Dajani and Gilbert (1978); Fielding et al. (1978, 1985a); Khan (1981); Barbour and Zerrillo (1982)
Exp / Veh h Op exp / Veh h Exp / Rev veh h	€ / veh h	McCrosson (1978); Fielding et al. (1978, 1985a); Giuliano (1981); Barbour and Zerrillo (1982); Kirby and Miller (1982); Vaziri and Deacon (1984); Talley (1988); Fielding

Op exp / Rev veh h		and Hanson (1988); Lee (1989)
Op lab exp / Veh km	€ / veh km	Holec et al. (1980), that compute it from a series of secondary indicators
Non-lab exp / Veh km	€ / veh km	Mackie and Nash (1982)
Op exp / Seat km	€ / seat km	Fielding et al. (1978); Keck et al. (1980)
Op exp / Seat h	€ / seat h	Keck et al. (1980)
Lab exp / Rev veh h	€ / veh h	Fielding et al. (1985a)
Lab exp / Seat km	€ / seat km	Dajani and Gilbert (1978)
Tot exp / Veh km	€ / veh km	Mackie and Nash (1982)
Tot exp / Seat km	€ / seat km	Lem et al. (1994). Costs include capital investments
Tot exp / Seat h	€ / seat h	Lem et al. (1994). Costs include capital investments
Tot exp / Empl	€ / person	McCrosson (1978)
Veh km / Veh Rev veh km / Veh	veh km / veh	Allen and DiCesare (1976); Fielding et al. (1978, 1985a); Dajani and Gilbert (1978); Keck et al. (1980); Mundle and Cherwony (1980); Barbour and Zerrillo (1982); Vaziri and Deacon (1984)
Veh h / Veh Rev veh h / Veh	veh h / veh	Fielding et al. (1978, 1985a); Keck et al. (1980); Fielding and Lyons (1981); Barbour and Zerrillo (1982); Talley (1988)
Veh km / Empl Rev veh km / Empl	veh km / person	McCrosson (1978); Fielding et al. (1978, 1985a); Mundle and Cherwony (1980); Khan (1981); Mackie and Nash (1982); Yeh et al. (2000)
Veh h / Empl Rev veh km / Empl	veh h / person	Fielding et al. (1978, 1985a); Fielding and Lyons (1981); Giuliano (1981); Vaziri and Deacon (1984); Fielding and Hanson (1988); Talley (1988)
Empl / Veh	person / veh	Mundle and Cherwony (1980)
Veh km / Empl h	veh km / person h	Barbour and Zerrillo (1982)
Veh h / Empl h	veh h / person h	Mundle and Cherwony (1980); Barbour and Zerrillo (1982); Lee (1989)
Rev seat km / Empl h	seat km / person h	Keck et al. (1980)
Rev seat h / Empl h	seat h / person h	Keck et al. (1980)
Veh km / Fuel Rev veh km / Fuel	veh km / liter	Fielding et al. (1978) consider energy rather than fuel consumption; Fielding and Lyons (1981); Fielding et al. (1985a)
Veh h / Fuel Rev veh km / Fuel	veh h / liter	Vaziri and Deacon (1984); Fielding et al. (1978) consider energy rather than fuel consumption
Seat km / Rev Veh	seat km / veh	Vaziri and Deacon (1984)
Platform h / Paid h	h / h	Allen and DiCesare (1976). Measures crew scheduling efficiency.
Veh h / Sched Veh h	veh h / veh h	Allen and DiCesare (1976). Measures fleet scheduling efficiency.
Rev veh h / Veh h	veh h / veh h	Giuliano (1981); Lee (1989)
2. Service use related to input		
Exp / Pax Op exp / Pax	€ / person	Fielding et al. (1978); Vaziri and Deacon (1984); Fielding et al. (1985a); McCrosson (1978) and Kirby and Miller (1982) consider total expenses
Exp / Pax km Op exp / Pax km	€ / person km	Fielding et al. (1978, 1985a); Dajani and Gilbert (1978); Keck et al. (1980); Mackie and Nash (1982); Vaziri and Deacon (1984); Kirby and Miller (1982) and Barbour and Zerrillo (1982) consider total expenses and Lem et al. (1994) also include capital costs
Deficit / Pax	€ / person	Talley and Becker (1982); Reciprocal of the shadow price of the service, as shown by Talley (1988)
Deficit / Pax km	€ / person km	Keck et al. (1980); Barbour and Zerrillo (1982)
Subs / Pax	€ / person	Kirby and Miller (1982); Fielding et al. (1985a)
Subs / Pax km	€ / person km	Kirby and Miller (1982)
Subs / Pax rev	€ / €	Fielding et al. (1985a)

Subs / Rev veh h	€ / veh h	Fielding et al. (1985a)
Rev / Veh op cost	€ / €	Barbour and Zerrillo (1982); Vaziri and Deacon (1984); Fielding and Hanson (1988)
Rev / Cost	€ / €	McCrosson (1978); Miller (1\980); Keck et al. (1980); Mackie and Nash (1982); Talley (1988); Lem et al. (1994) also consider capital costs
Pax rev / Op exp	€ / €	Vaziri and Deacon (1984)
Pax / Empl	person / person	Allen and DiCesare (1976); Yeh et al. (2000)
Pax / Empl h	person / person h	Keck et al. (1980); Barbour and Zerrillo (1982)
Pax km / Empl h	person km / person h	Keck et al. (1980); Barbour and Zerrillo (1982)
Pax / Fuel	person / liter	Fielding et al. (1985a); Talley (1988) considers energy rather than fuel consumption
Pax km / Fuel	person km / liter	Dajani and Gilbert (1978)
3. Intensity of service use		
Pax / Veh	person / veh	Fielding et al. (1978)
Pax / Veh km Pax / Rev veh km	person / veh km	McCrosson (1978); Fielding et al. (1978, 1985a); Austin and Stone (1980); Fielding and Lyons (1981); Vaziri and Deacon (1984); Lee (1989); Yeh et al. (2000)
Pax / Veh h Pax / Rev veh h	person / veh h	McCrosson (1978); Fielding et al. (1978, 1985a); Austin and Stone (1980); Miller (1980); Fielding and Lyons (1981); Barbour and Zerrillo (1982); Fielding and Hanson (1988); Talley (1988); Lee (1989)
Pax / Seat km	person / seat km	Allen and DiCesare (1976)
Pax / Seat h	person / seat h	Keck et al. (1980)
Pax / Lines length	person / km	Dajani and Gilbert (1978); Vaziri and Deacon (1984); Musso and Vuchic (1988)
Pax km / Veh km	person km / veh km	Khan (1981); Barbour and Zerrillo (1982); Mackie and Nash (1982); Lee (1989)
Pax km / Veh h	person km / veh h	Barbour and Zerrillo (1982)
Pax km / Seat km	person km / seat km	Vaziri and Deacon (1984); Tulkens and Wunsch (1994); Keck et al. (1980); Musso and Vuchic (1988); Lee (1989); Lem et al. (1994)
Pax km / Seat h	person km / seat h	Keck et al. (1980); Lem et al. (1994)
Pax km / Lines length	person km / km	Allen and DiCesare (1976)
4. Relative service dimension		
Exp / Pop Op exp / Pop	€ / person	Vaziri and Deacon (1984); Kirby and Miller (1982) consider total and local system costs
Op subs / Pop	€ / person	Kirby and Miller (1982); Vaziri and Deacon (1984)
Veh / Pop	veh / person	Vaziri and Deacon (1984)
Empl / Pop	person / person	Vaziri and Deacon (1984)
5. Service coverage		
Seat km / Pop	seat km / person	Vaziri and Deacon (1984)
Rev veh km / Pop	veh km / person	Vaziri and Deacon (1984)
Lines length / Area	km / km^2	Allen and DiCesare (1976); Vaziri and Deacon (1984); Musso and Vuchic (1988)
Lines length / Pop	km / person	Musso and Vuchic (1988)
Veh km / Area	veh km / km^2	Allen and DiCesare (1976)
Lines length * WD / Area	km * km / km^2	Allen and DiCesare (1976); WD = walking distance
Covered area / Area	km^2 / km^2	Musso and Vuchic (1988)
Covered pop / Pop	person / person	Allen and DiCesare (1976); Fielding et al. (1978); Dajani and Gilbert (1978); Fielding and Lyons (1981); Talley (1988)
6. Market penetration		
Pax / Area	person / km^2	Allen and DiCesare (1976); Fielding et al. (1978); Talley (1988)
Pax / Pop	person / person	Allen and DiCesare (1976); Khan (1981); Fielding and

		Lyons (1981); Kirby and Miller (1982); Vaziri and Deacon (1984); Musso and Vuchic (1988)
Pax km / Area	person km / km^2	Vaziri and Deacon (1984)
Pax km / Pop	person km / person	Dajani and Gilbert (1978)
Trip length / EAR	km / km	Vaziri and Deacon (1984); EAR = equivalent area radius
7. Revenues generation		
Rev / Veh Pax Rev / Veh	€ / veh	Vaziri and Deacon (1984); Fielding et al. (1985a)
Rev / Veh km	€ / veh km	McCrosson (1978)
Pax rev / Rev veh h	€ / veh h	Kirby and Miller (1982); Fielding et al. (1985a); Lee (1989)
Rev / Empl	€ / person	Mackie and Nash (1982)
Rev / Pax Pax rev / Pax	€ / person	McCrosson (1978); Lee (1989); Kirby and Miller (1982)
Rev / Pax km Pax rev / Pax km	€ / person km	Kirby and Miller (1982); Keck et al. (1980) and Barbour and Zerrillo (1982) consider revenues as farebox plus local subsidies only (therefore excluding those from U.S. federal programs)
8. Externalities		
Cost / User benefit	€ / €	Kirby and Miller (1982)
User benefit / Pop	€ / €	Kirby and Miller (1982)
Cost / Veh km reduced	€ / veh km	Kirby and Miller (1982)
Veh km reduced / Pop	veh km / person	Kirby and Miller (1982)
Rev veh km / Call	veh km / call	Vaziri and Deacon (1984); Fielding et al. (1985a); call = number of calls of the vehicle to the operational centre for a service disruption
Acc / Rev veh km	event / veh km	Allen and DiCesare (1976); Dajani and Gilbert (1978); Fielding and Lyons (1981); Vaziri and Deacon (1984); Fielding et al. (1985a); Fielding and Hanson (1988)
Acc / Rev veh h	event / veh h	Fielding et al. (1985a)

By looking at the indicators in the table, we conclude that the following four aspects of the operation of the service are the most considered ones when evaluating a public transport system: (1) the resources being used to produce the service, (2) the quantities of service being produced, (3) the quantities of service produced that are actually consumed (i.e. the patronage levels) and (4) the potential dimension of the market for the service. Other considered aspects are (5) the service revenues, that in the transport sector are usually only a fraction of the total revenues, and (6) the service externalities, whose importance is apparent in our framework, where we postulate that they must be fully considered in the evaluation of public transport. An indicator usually puts in relation two of these six aspects. Table 2 therefore clusters the reviewed indicators according to the following groups, where indicators pertaining to the same group consider the same aspects while using different measures:

1. *Technical efficiency*, to relate the quantities of produced service with the resources being used.

2. *Service use related to input*, to relate the resources being used with the patronage.

3. *Intensity of use of the service*, to relate the quantities of produced service with the patronage.

4. *Relative service dimension*, to relate the resources being used with the dimension of the potential market.

5. *Service coverage*, to relate the produced service with the dimension of the potential market.

6. *Market penetration*, to relate the patronage with the dimension of the potential market.

7. *Revenues generation*, to relate the revenues with consumed resources, produced service and patronage.

8. *Externalities*, to relate some key effects of the system operations with consumed resources, produced service and patronage. According to the literature, the table only considers benefits for users and environmental benefits in terms of decreasing motorized trips on the positive side, and accidents as negative externalities. This latter is the most frequently considered external effect in the reviewed evaluation exercises.

It can be seen that efficiency and effectiveness concepts, as defined in the preceding section, are somewhat transversal to this classification. In particular, efficiency measures are more related with those indicators belonging to groups 1, 2, 3, 6 and 7 that consider the managerial perspective. Effectiveness indicators should instead reflect the point of view of the general interest, and are therefore primarily to be found in groups 4, 5, 6 and 8.

We note that such distinction does not respect in some cases the classification originally proposed by the authors. This derives from our theoretical framework laid out in the previous section, that might differ from the one of these papers. For example, many of the authors listed in the third group consider the corresponding measures as effectiveness indicators, consistently with the definitions of effectiveness that are reported in table 1, whereas we would say these are efficiency measures. Moreover, Gleason and Barnum (1982) point out that there has always been some confusion between efficiency and effectiveness concepts and the related indicators. In general terms, many so-called effectiveness indicators are ratios containing amounts of resources or inputs, thus becoming also indicators of efficiency.

Another source of inconsistency is due to the fact that the point of view of service users is sometimes not distinguished from that of the general population. For example, Hensher and Prioni (2002) note that maximising the service in terms of passenger-km was considered by British rail a good proxy of social welfare up to the Seventies. However it should be noted that the point of view and the interests of (actual or potential) public transport users is not necessarily that of the general public, particularly in those countries or urban agglomerations where public transport is used or practically usable by a small minority of people. Considering the general interest, good public transport could for example be desirable if it can divert trips from private motorized transport means thus reducing congestion and pollution, less so if it attracts trips previously made by bike or if it encourages urban sprawl.

Quality indicators

The above review does not exhaust the range of indicators that have been proposed to evaluate public transport. In more recent years, modal diversion from cars to public transport has more and more been considered as a key tool to improve transport systems (European Commission, 2001). More recent evaluation methods tend therefore to consider levels of service and quality aspects beyond economic efficiency, since public transport should be able to offer a service that is competing with that of private cars. This evolution is in relation with the increased attention paid to the point of view of the transport users. Marketing aspects are therefore also considered when evaluating the relative performances of different transport modes. More recent literature has therefore proposed a wealth of indicators related to several aspects of service quality.

We present in the following table 3 a selection of these quality indicators. This table has the same structure of the preceding table 2. Since our focus is on effectiveness evaluations, we did not consider those works related to service quality and customer satisfaction that do not present a synthetic indicator (such as psychometric techniques to quantify comfort-related constructs), nor works dealing with decision support tools for the planning and design of the service. For the same reason, studies on accessibility not specifically focussed on public transport systems have also been excluded, together with those papers that study qualitative criteria (cleanliness, staff courtesy etc.) through analytic methods different from indicators, such as scaling and rating exercises.

Table 3 – Quality indicators for public transport

Quality Indicator	*Measurement Unit*	*References and Notes*
1. On-time performance		
R = f(E, L, h)	Min	Alter (1976) reliability, tabulated as a function of the % of transit earlier than 1 min (E), the % of transit later than 3 min (L) and the headway (h)
R = 1/ S(tt)	1 / min	Sterman and Schofer (1976) reliability. S(tt) = standard deviation of travel times
Pax on time / Pax	person / person (%)	Buneman (1984); Seco and Gonçalves (2007) report 12 other studies using it
Pax minutes delay / Pax	person min / person	Buneman (1984)
$R = 1 - \frac{2\sum_{r=1}^{n}\left(h_r - \bar{h}\right)r}{n^2\bar{h}}$	-	Henderson et al. (1991) regularity index. h = series of n headways with index r from smallest (r=1) to largest (r=n)
$W = \frac{1}{1 - C_v^{\,2}} = \frac{1}{EIS}$	-	Henderson et al. (1991) passenger wait index. C_v = coefficient of variation of headways. Its reciprocal is the evenness index based on stops (EIS), according to Chen et al. (2009)
OTP	min / min (%)	On-time performance used in New York (Nakanishi, 1997). OTP = % of trips with less than 5 min delay
Service regularity	min / min (%)	Service regularity in New York (Nakanishi, 1997), as the % of actual headways within a given band width of the scheduled one
$R = \frac{\sum_{k=1}^{H} k \cdot p(k)}{H}$	-	Camus et al. (2005) weighted delay index. H = scheduled headway ranging from 1 to k minutes, p(k) = observed probability for delay of k minutes
$DIS = p\{H_s - H_0 \in [a, b]\}$	-	Chen et al. (2009) deviation index based on stops. H_s = headway at stop s, H_0 = headway at terminal, [a, b] = time

		interval
$PIR = p\{t_{run} - t_{sch} \in [a, b]\}$	-	Chen et al. (2009) punctuality index based on routes. t_{run} = actual running time of one route, t_{sch} = scheduled running time of one route, [a, b] = time interval
2. Relative extent of the service		
Service frequency	veh / h	Botzow (1974); Seco and Gonçalves (2007) report 9 other studies using it; Alter (1976) jointly considers population density
Weighted peak frequency	veh / h	Musso and Vuchic (1988); weights based on stops
ITSA	-	Henk and Hubbard (1996) index of transit service availability, average of the route km per km^2, the veh km per route km and the seat km per capita
TLOS	person h / person h (%)	Ryus et al. (2000) level of service based on availability. TLOS = percent of persons hour for which transit is available
A = T / Q	trips / person	Polzin et al. (2002) accessibility index (trips per capita for which transit is available). T = total daily trips for which transit is available, Q = equivalent population considering also workplaces
% jobs served by transit	%	TCRP (2003a) accessibility
% resid. served by transit	%	TCRP (2003a) accessibility
TTR	-	Yin et al. (2004) system wide reliability. TTR = probability that the average travel time is below a given threshold
SR	-	Yin et al. (2004) schedule reliability. SR = probability that the average headway is below a given threshold
WTR	-	Yin et al. (2004) wait time reliability. WTR = probability that the average wait time is below a given threshold
3. Relative performance of the service*		
A = TTB / TTA	min / min	Alter (1976) relative travel time. TTA = travel time by car; TTB = travel time by bus
$R = TT_{hm} / tt_{hm}$	min / min	Lam and Schuler (1982) connectivity index. TT_{hm} = harmonic mean of reference trip lengths; tt_{hm} = harmonic mean of trip lengths by public transport
A = 2*TTB / (TTA + TTB)	min / min	TCRP (2003a) accessibility index. TTA = travel time by car; TTB = travel time by bus
TSI = WTA / WTT	min / min	Fu and Xin (2007) transit service indicator. WTA = door-to-door car travel time weighted with walking components; WTT = door-to-door transit travel time weighted with walking components
$F = C_{PT} / C_C$	€ / €	Seco and Gonçalves (2007) relative fare level. C_{PT} = out-of-pocket public transport cost; C_C = out-of-pocket car cost
4. Other quality aspects		
Commercial speed	km / h	Botzow (1974); Mundle and Cherwony (1980); Seco and Gonçalves (2007) report 11 other studies using it
Weighted operational speed	km / h	Musso and Vuchic (1988); weights based on veh km per time
Walking time to/from stops	min	Alter (1976) basic accessibility
D = f(NT, WT)	-	Alter (1976) directness, tabulated on the basis of the number of transfers NT (0, 1, 2, more than 2) and the wait time WT
Transfer Pax / Pax	person / person (%)	Allen and DiCesare (1976); Seco and Gonçalves (2007) report 6 other studies using it
O = P / S	person / seat	Vehicle occupancy = Persons onboard / Available seats; Alter (1976) also considers standing places and the surface per pax

* A review of quality indicators concerning service performances is also reported in Bhat et al. (2005)

Indicators from table 3 have been grouped according to the following four main service evaluation dimensions:

1. From our review, it is apparent that the most frequently investigated issue is related to the regularity of the schedule.

2. The size of the system in relation with several factors, such as the frequency and the size of the service area, is the second most "popular" topic being considered to study the service quality.

3. Researchers have also focussed on the performance of the system in relation with its major competitor, namely private cars, consistently with the emphasis that was given in the last decade to modal diversion issues.

4. Several other quality aspects are also considered, such as service speed, walking times and transfers. However, as we already mentioned, these latter aspects are often investigated through methods that do not rely on indicators, so that their importance is not fully highlighted in this review.

Although the focus is here on quality and levels of service from the user viewpoint, we note that several indicators from this table could be related to service effectiveness. For example, agencies could be willing to fund transport services in order to improve accessibility by public transport to different territorial functions, according to those definitions of accessibility and availability that are related to indicators pertaining to the second group. This makes such indicators also relevant to study effectiveness. Even more importantly, measures relating the performance of public transport with that of other transport modes (primarily cars) are quality measures for public transport users, but they can also give an indication on the extent to which public transport can divert trips from more polluting modes. Indicators of the third group needs therefore also to be considered to evaluate effectiveness if modal diversion is among the policy goals.

As a concluding remark, we note from table 3 that researchers adopt different definitions for commonly used terms such as reliability and accessibility, that can in fact be found to define indicators that are very different each other and scattered in different sections of the table. We do not analyse thoroughly this issue since it is not the focus of the paper, however it is clear that also this lack of consistency in defining these key concepts can contribute in jeopardizing an evaluation exercise.

Relevant indicators either from manuals or outside the public transport field

Beyond the above reviewed more specialised scientific literature in the public transport field, on more applicative grounds we note that the U.S. Transit Cooperative Research Program details a methodology that allows for a full evaluation through the sequential use of different sets of indicators (TCRP, 2003a). A series of selection menus is proposed, through which the reader can find the right set of indicators to use according to the objectives of the evaluation exercise. Several dozens of indicators are therefore proposed; the most relevant ones are included in the preceding table 3, since they were mostly taken from the scientific literature. Another relevant report from the same research program focuses on quality of service aspects, by proposing more elaborate analytical methodologies beyond the use of indicators (TCRP, 2003b).

Evaluation exercises in the transport sector are of course not limited to public transport. Therefore, effectiveness indicators to monitor the compliance of a transport system to a policy goal can be found also from other sources, particularly concerning environmental issues. In particular, the Transport and Environmental Reporting Mechanism (TERM) of the European Environmental Agency explicitly takes the point of view of the social community that is more concerned with the external impacts. A set of effectiveness indicators from a list of 40 items are annually monitored, thus offering an overview of the environmental impact trends in the transport sector (EEA, 2009). Those indicators are drawn from national data sources, ranging from land use (land occupancy by infrastructures, habitat fragmentation) to aggregate emissions to traffic measures. From this same perspective, several EU projects have investigated the role of indicators in transport planning and decision making processes (e.g. QUATTRO, 1998; May, 2005; REFIT, 2008; Calderón et al., 2009) and transport indicators selection methods according to the evaluation framework have also been proposed (e.g. Castillo and Pitfield, 2010). Even if such literature is surely related to our perspective, we do not extensively review it here, since its scope is much wider compared to ours, that is instead restricted to the evaluation of the operations of public transport systems.

Concluding remarks on the review of indicators

According to the framework set up in the second section, we believe that decision makers need analysis tools that can effectively distinguish among the different perspectives of the actors of a transit system, namely funding agencies, service operators and service users. This seems to have been only partially achieved when considering the state of the art related to the development of performance indicators, where these different perspectives are sometimes more mingled together. Past research efforts rather focussed in bringing into evidence all the different aspects that should be considered, therefore proposing a wide variety of indicators. Within such body of work, it seems to us that efficiency and service quality issues have been properly addressed, whereas indicators that are relevant to the public interest viewpoint seem to be more dispersed across the different groups of measures that are reported in tables 2 and 3. Therefore, in the following section we would like to propose a preliminary and more coherent set of public transport effectiveness indicators that builds on the above reviewed works, in order to provide decision makers with a decision support tool that is more in line with their need of matching public transport services operations with the more general transport policy goals.

PROPOSAL FOR A SET OF EFFECTIVENESS INDICATORS FOR PUBLIC TRANSPORT

Indicators evaluation dimensions and wished properties

Miller and Kirby (1984) note that public transport objectives are very wide, however related benefits can essentially be determined by measuring the impacts of the system on travel behaviours in terms of (1) mobility rates, (2) trips schedules and patterns, and (3) trip purposes and modes being used. In other words, one should be able to evaluate even more indirect influences of transit services by always looking at transport-related quantities. We believe that this is a good starting point for developing a performance assessment system that considers the general interest's point of view. However these authors do not explicitly define efficiency and effectiveness in their study in order to separate the different perspectives, only referring to the concept of performance evaluation. In the following we explicitly deal with effectiveness, according to the framework of the preceding sections. Therefore, we exemplify the evaluation procedure by considering the following policy effectiveness dimensions, that match the most common socially relevant objectives that public transport should reach:

1. *Territorial accessibility* for those exclusively using the public transport service

2. *Relative performances* compared with private motorized transport means

3. *Environmental impacts* of the whole transport system (all modes jointly considered)

These three dimensions were primarily considered because they are the most frequently investigated ones in the literature. It can also be noted that, taken together, they allow the analyst to indirectly consider many other benefits of public transport. This is possible because the indicators that we are going to present are mainly output indicators, i.e. looking at the direct results of transit operations. Each of these three main outputs can however be linked to different outcomes, i.e. more far-reaching implications, that in turn allow the analyst to indirectly consider also other relevant aspects through our methodology. After having presented the three indicators related to the above dimensions, we therefore consider how these could form the basis of a more comprehensive evaluation exercise that encompasses several other aspects.

The indicators that we propose in the following subsections are mainly based on the literature review that has been previously completed. Of course, other alternative or additional indicators could be foreseen as well. In the following we just want to give examples of indicators that are consistent with our evaluation framework.

Several authors (e.g. Morris et al. 1979; Allen and Grimm, 1980; Miller, 1980; Mackie and Nash, 1982; Giannopoulos, 1989; Meyer and Miller, 2001; Seco and Gonçalves, 2007) list the desirable characteristics of "ideal" performance indicators for transit, mainly from the point of view of the transport practitioner. Given the preliminary nature of the present work, in the following we are more keen to adopt the point of view of the researcher, where some elements such as the computational easiness are less important. With this in mind, we think that the following properties of an effectiveness indicator are desirable:

1. they should be strictly related to the objective of the evaluation exercise;

2. data needed to compute them should be available in the study context;

3. analytical procedures to compute them should fit the data (for example, adequately treating non-metric or qualitative information);

4. they should consider elements that can be directly or indirectly influenced by the subject on which the evaluation is carried (public transport operator, transit agency, policy maker), so that they are effectively responsive to those actions that can be taken by the subject itself;

5. they should be robust, so that they can be used at different points in time to monitor trends, but also across different services to draw comparisons;

6. they should have some intuitive meaning and they should be easy to interpret; in particular they should be unambiguous, so that their variation in a given direction can be safely considered as good or bad.

In the following we will evaluate to which extent the proposed indicators match those properties.

Territorial accessibility indicator

Social inclusion and the reduction of inequalities have been among the chief objectives of most public transport funding programs, at least since when cars have conquered a predominant role in urban mobility, with the related risk of discriminating those that for any reason cannot use it. In order to monitor how well the system is reaching this objective, several different indicators have been proposed to measure accessibility, as shown in the preceding tables. When looking at the “best” index according to our framework, we believe that the following elements should be considered:

1. Due to the derived utility of the travelling activity, accessibility to territorial opportunities through the public transport system should ideally be considered, rather than the accessibility to the system itself (i.e. “accessibility through transit”, and not “accessibility to transit”), in order to meet the above proprieties 1 and 6 among those listed in the preceding subsection. However a much richer dataset is obviously needed for such analysis, typically coming from a household travel survey.

2. Consistently with property 6, an accessibility indicator should be related to the actual travel patterns, rather than to the mere possibility of using the service by a group of people, independently on where they have to go. This latter measure is more related to the idea of service availability. However, accessibility and availability concepts are not always clearly distinguished in the literature.

3. A good accessibility indicator should focus more on the process of the service production than on its outcome in terms of behavioural response, as Morris et al. (1979) point out. Behaviours can in fact be influenced by a wide range of elements that are beyond the control of the service provider. This is consistent with the previously reported fourth property, i.e. the importance of only considering elements that can be affected by the service provider when running an evaluation activity.

4. Trip-based measures are more accurate than passenger-based ones, since onboard passenger counts might not consider the fact that passenger can use more than one transit line to complete their journey. As a consequence, systems with higher rate of corresponding trips would paradoxically look like offering better accessibility.

5. We should only consider accessibility through transit, since any relative performance measure with other competing modes such as cars are covered by the indicator proposed in the following section.

By looking at the different accessibility measures that have been proposed in the literature, we can realise that several of them do not meet one or more of the above criteria. Among the remainder, it seems to us that a good indicator has been proposed by Polzin et al. (2002). Their methodology allows to determine the portion of trips that could be served by transit considering both spatial and temporal availability through traffic analysis zones (TAZ) information (zone population and trips), route buffers and temporal distribution of travel demand. This approach could even be improved when microdata from a travel survey are available in the study area, thus studying accessibility at the trip rather than at the zone level. In such case, an indicator of accessibility could be defined as follows:

1. For each trip in the dataset, define which trip end is prevalent in determining the actual trip timing. For example, for home-based trips this would most likely be the non-home trip end. If the activities performed at different locations are available this step is easy, whereas if only “trip purposes” are known the task could be more tricky and some heuristic rules would probably be needed.

2. Compute suitable time windows of the trip ends. To do so, the analyst should fix some upper bounds to the maximum wait or advance time at bus stops (both at the beginning of the journey and for eventual correspondences) and the maximum ride time (not on the basis of the ride times of competing modes such as cars, as we explain below, but rather of non motorized means such as feet or bike). The time window related to the above defined prevailing trip end should be computed first. We do not detail the procedure, since it is well known from the Demand Responsive Transport scheduling field: we refer the interested reader for example to Diana and Dessouky (2004).

3. Determine if the service can serve the trip within such time windows and without requiring walking to/from bus stops above a given thresholds.

4. The accessibility indicator would then be the ratio between trips that can be served (or are already served) by transit and the total trips in the dataset.

The above indicator is well fitting properties 1, 2 (when data from a travel survey are available), 3, 5 and 6 of the list in the previous section, and is rather robust also on the fourth one, in particular when schedules are not affected by car congestion.
We note that a similar methodology has been applied in the context of studying the potential for modal diversion of public transport (Massot et al., 2006), although in that work the performances of competing modes (namely, cars) have been considered, that would lead to an indicator not respecting the fourth property. Within our public transport evaluation context, where the focus is not in predicting the actual modal diversion potential since this involves a behavioural change issue, such aspect is rather considered in the following subsection.

Relative performances indicator

One central issue on a transport policy viewpoint is to divert trips from private motorized transport means. In order to reach this goal, it is clear that one should monitor the relative performance of public transport related to the other modes, rather than simply looking at how good the public transport system is in itself. The resulting indicators would probably not fit the fourth above property if we consider the point of view of the transport firm, and they should therefore not be used to evaluate the operators. However, from a policy effectiveness viewpoint, transport decision makers have also the opportunity to influence the performances of private cars through several measures, for example concerning traffic regulations. Therefore, the following indicator can evaluate the effectiveness of a broader set of measures aimed at achieving modal diversion, beyond those directly related to public transport provision. Unlike accessibility measures discussed in the preceding subsection, the following indicators consider data from different competing modes.
To evaluate the relative performances of the service, we propose a modification of the Lam and Schuler (1982) connectivity indicator that we presented in table 3. These authors measure connectivity on a representative set of *n* trips as the ratio between the harmonic mean of “reference” trip travel times, that are a sort of lower bound that could be obtained if the service were fully developed, and the harmonic mean of trip travel times done by public transport in its real configuration.
Connectivity in their work can essentially be seen as the joint consideration of territorial accessibility, as discussed in the preceding subsection, and level of service provided. It is therefore possible to more easily draw comparisons across systems that operate in different regions, where territorial configurations (e.g. corridor-like versus more dispersed urban areas) generally lead to biased results, as discussed in their work. Another nice feature of this index is the presence of two of the above mentioned desirable features for accessibility indicators, namely process orientation and territorial rather than service accessibility. On the mathematical side, harmonic rather than arithmetic means are considered, thus lessening the influence of higher performance gaps and longer trips. Both these effects seem appropriate in our case, leading to an increase of the sensitivity of the evaluation method whenever there is not a mode that clearly outperforms the others (as demand levels do not

linearly vary with offer characteristics) and whenever shorter trips are considered (since these latter usually represent a priority target of modal diversion in urban areas).
To sum up, we believe that such connectivity indicator has a great potential to assess a public transport system on a policy viewpoint. However we believe that trip travel times could possibly be replaced with generalized travel costs in a modelling framework. Another improvement would consist in considering the reference values (indicated as “TT” in the table) not on the basis of graph theory, but rather considering the best performing travel mode for each trip, i.e. the mode that has the minimum generalized travel cost. Therefore, the new index *R* would be the following:

$$R = \frac{\sum_{i=1}^{n} \frac{1}{c_i}}{\sum_{i=1}^{n} \frac{1}{C_i}}$$

where *n* is the number of trips being considered, c_i is the generalised travel cost of trip *i* done by public transport and C_i is the generalised travel cost of trip *i* made with the “best” mode. This index ranges from 0 to 1, higher values indicating more competitive public transport systems. The modification of this index that we propose would allow the analyst an immediate feeling of the relative performance of the public transport service against the other competing modes, allowing to quickly appreciate its potential to attract riders. Focussing on modal diversion from cars, generalized costs of car trips could be considered, rather than those of the best mode, in order to compute C_is.
Reviewing this indicator in light of the six above defined desirable properties, let alone the above discussed critical point concerning the fourth one that however seems justified considering the goals of the evaluation, a simulation model is probably needed to obtain C_is, thus limiting the applicability of the method to situations where such model is available.

Environmental impacts indicator

Transport systems affect the environment in several different ways. However, within the limits of the assessment of public transport operations, we believe that the chief benefit is the net variation in the vehicle kilometres globally travelled (VKT) in a given territory, as suggested for example by Kirby and Miller (1982). This could then be translated in a corresponding variation of emissions of pollutants and greenhouse gases through an emission model. It is clear that the corresponding indicator is more problematic concerning the fourth property in the above list, since VKT reductions depend on behavioural responses of individuals that are outside the control of decision makers. On the other hand, this indicator can indirectly capture qualitative elements related to the effectiveness of marketing and promotion actions of public transport that are not considered in the previously discussed indicators, since they are difficult to include in a generalised transport cost function. Concerning the other properties, this indicator presents the same benefits and drawbacks of the previous one on the relative performances of the different means. Their joint use seems therefore appropriate.

The computation of such indicator could be done by considering the demand actually being served by the service under evaluation and estimating the proportion of trips that would still be made through private motorized means if the service did not exist. For such trips, the analyst would therefore run a traffic assignment model to estimate the VKT generated. The difference between such quantity and the VKT travelled by the service represents an environmental impact indicator in absolute terms. Such absolute quantity could therefore be related to the total VKT being travelled in the study area to give a clearer indication of the real impact of the system.

Considering additional evaluation aspects and different frameworks

The three above ratio indexes cover three key elements in the evaluation of any public transport system. Beyond that, one can notice that they can also provide indications on how the system is performing from a much wider range of viewpoints. For example, accessibility is linked to issues such as social inclusion and livability of the urban environment. Modal diversion from less safe to safer modes, as captured by the second indicator, and variations in VKT, that are the object of the third one, are two outputs of the system operations that can be a proxy of several outcomes such as safety improvements or congestion decrease. To sum up, jointly considering these three indicators together can be seen as a reasonably complete evaluation exercise, since they can give useful hints also on aspects different from those that we directly considered.

These indicators could be used both to monitor the evolution of the effectiveness of a system (temporal evaluation) and to compare the effectiveness of different systems (spatial evaluation). In this latter case, particular care must be taken in order to draw comparisons that are meaningful, as discussed at the beginning of the second section. This can be achieved by adopting a method to define peer groups of transit providers within which such comparative assessment can be safely carried out. Such methods are presented in the papers mentioned in the second section.

CONCLUSIONS AND FUTURE WORK

The present paper has reviewed public transport indicators that have been proposed in the literature. The abundance of references is a clear indication of the advantages of the use of synthetic indicators to perform such evaluation exercise. Most of the indicators that we have found have a simple formulation, an intuitive meaning and are computationally quite easy to calculate. This allows for their widespread use also beyond the research community, although the required data might not be immediately available in some cases. However when limiting the analysis to the consideration of one or more indicators, as it is customarily done, the researcher must pay close attention to the proper ambits of application of the indicator itself, since it is not infrequent to see evaluation exercises that dramatically aborted because of the improper use of some indicator.

Our research was focussed on the evaluation needs of public bodies that finance the operations of the service. The starting point of our work is that the point of view of the decision maker might not coincide with that of the actors of the system, i.e. public transport operators and customers. In the authors' opinion, effectiveness evaluations should therefore be referred only to the policy maker point of view, in order to avoid misunderstanding or even inconsistent results from the evaluation activity. Considering this framework, existing indicators in the literature have been classified and it has been concluded that an effectiveness analysis should consider indicators that do not necessarily coincide with those used to monitor the efficiency of the firm or the quality of the provided service.

The last part of the paper has been devoted to the proposal of a set of indicators that is consistent with the above framework. However by no way we postulate that the proposed indicators are the best ones in absolute terms. A solution that is always superior to every other is in fact simply impossible to find, because the desirable properties of transport indicators are in some cases conflicting, thus leading to a classical multicriteria problem. The present paper was in fact strongly focussed on the review of the indicators that have been proposed to evaluate public transport, in such a way that a researcher or practitioner can find the most appropriate ones in a given set of circumstances.

We finally point out that the analysis that we have conducted could also be extended beyond the study of performance indicators. Indicators can in fact generally give only a limited insight on the complexity of the transport matter, so that the researcher often risks to miss some important elements that could give a more complete picture of the situation. Many of these limitations could be overcome by embedding the use of public transport indicators in an economic analysis framework (Daraio and Diana, 2009).

REFERENCES

Allen, W.G.; DiCesare, F. (1976). Transit service evaluation: preliminary identification of variables characterizing level of service. Transportation Research Record, 606, 41-47.

Allen, W.G.; Grimm, L.G. (1980). Development and application of performance measures for a medium-sized transit system, Transportation Research Record, 746, 8-13.

Alter, C.A. (1976). Evaluation of public transit services: the level-of-service concept, Transportation Research Record, 606, 37-40.

Austin Jr., J.A.; Stone, T.J. (1980). The application of statistical control measures to transit performance indicators. Journal of Advanced Transportation, 14(3), 213-236.

Barbour, L.C.; Zerrillo, R.J. (1982). Transit performance in New York state. Transportation Research Record, 857, 18-25.

Bhat, C.R.; Guo, J.Y; Sen, S.; Weston, L. (2005). Measuring access to public transportation services: Review of customer-oriented transit performance measures and methods of transit submarket identification. Report No. FHWA/TX08/0-5178-1, Center for Transportation Research, The University of Texas at Austin. Available at http://www.caee.utexas.edu/prof/bhat/REPORTS/5178_R1_rev_Sept05.pdf - Accessed 19th January 2010.

Botzow, H. (1974). Level-of-service concept for evaluating public transport. Transportation Research Record, 519, 73-84.

Brons, M.; Nijkamp, P.; Pels, E. ; Rietveld, P. (2005). Efficiency of urban public transit: A meta analysis. Transportation, 32(1), 1-21

Buneman, K. (1984). Automated and passenger-based transit performance measures. Transportation Research Record, 992, 23-28

Calderón, E.J.; Pronello, C.; Goger, T. (2009). Integrated assessment of environmental impact of traffic and transport infrastructure. COST Office and Universidad Politécnica de Madrid – Servicio de Publicaciones, Madrid, 405pp, ISBN 978-84-7493-401-4.

Camus, R.; Longo, G.; Macorini, C. (2005). Estimation of transit reliability level-of-service based on automatic vehicle location data. Transportation Research Record, 1927, 277-286.

Castillo, H.; Pitfield, D.E. (2010). ELASTIC – A methodological framework for identifying and selecting sustainable transport indicators. Transportation Research D, 15(4), 179-188.

Chen, X.; Yu, L.; Zhang, Y.; Guo, J. (2009). Analyzing urban bus service reliability at the stop, route, and network levels, Transportation Research A, 43(8), 722-734.

Dajani, J.S.; Gilbert, G. (1978). Measuring the performance of transit systems. Transportation Planning and Technology, 4(2), 97-103.

Daraio, C.; Diana, M. (2009). Efficiency and effectiveness in the public transport sector: a critical review with directions for future research. Presentation at the XI European Workshop on Efficiency and Productivity Analysis, Pisa (Italy), 23-26 June.

De Borger, B.; Kerstens K.; Costa, A. (2002). Public transit performance: What does one learn from frontier studies?. Transport Reviews, 22(1), 1-38.

Diana, M.; Dessouky, M.M. (2004). A new regret insertion heuristic for solving large-scale dial-a-ride problems with time windows. Transportation Research B, 38(6), 539-557.

EEA – European Environmental Agency (2009). Transport at a crossroads – TERM 2008: indicators tracking transport and environment in the European Union. Luxembourg, 52pp., ISBN 978-92-9167-991-1.

European Commission (2001). White Paper ''European transport policy for 2010: time to decide''. Luxembourg, 119pp., ISBN 92-894-0341-1.

Farrell, M.J. (1957). The measurement of productive efficiency. Journal of the Royal Statistical Society, Series A (General), 120(3), 253-290.

Fielding, G.J.; Glauthier, R.E.; Lave, C.A. (1978). Performance indicators for transit management. Transportation, 7(4), 365-379.

Fielding, G.J.; Lyons, W.M. (1981). Performance evaluation for discretionary grant transit programs. Transportation Research Record, 797, 34-40.

Fielding, G.J.; Babitsky, T.T.; Brenner, M.E. (1985a). Performance evaluation for bus transit. Transportation Research A, 19A(1), 73-82.

Fielding, G.J.; Brenner, M.E.; Faust, K. (1985b). Typology for bus transit. Transportation Research A, 19A(3), 269-278.

Fielding, G.J.; Hanson, L. (1988). Determinants of superior performance in public transit: Research opportunities using Section 15 data. Transportation Research Record, 1165, 94-98.

Fu, L.; Xin, Y. (2007). A new performance index for evaluating transit quality of service. Journal of Public Transportation, 10(3), 47-69.

Giannopoulos, G.A. (1989). Bus planning and operation in urban areas: A practical guide. Avebury, Aldershot, UK, 370pp.

Giuliano, G. (1981). Effect of environmental factors on the efficiency of public transit service. Transportation Research Record, 797, 11-16.

Gleason, J.M.; Barnum, D.T. (1982). Toward valid measures of public sector productivity: Performance measures in urban transit. Management Science, 28(4), 379-386.

Henk, R.H.; Hubbard, S.M. (1996). Developing an index of transit service availability. Transportation Research Record, 1521, 12-19.

Hensher, D.A.; Prioni, P. (2002). A service quality index for area-wide contract performance assessment. Journal of Transport Economics and Policy, 36(1), 93-113.

Holec Jr, J.M.; Schwager, D.S.; Fandialan, A. (1980). Use of federal Section 15 data in transit performance evaluation: The Michigan program. Transportation Research Record, 746, 36-38.

Keck, C.A.; Zerrillo, R.J.; Schneider, N.R. (1980). Development of multimodal performance measures for transit systems in New York state. Transportation, 9(4), 355-367.

Kelley, W.J.; Rutherford, G.S. (1982). Transit system performance evaluation methodology – An interim draft report. Report No. WA-RD-51.1, Department of Civil Engineering, University of Washington, 87pp. Available at http://www.wsdot.wa.gov/research/reports/fullreports/051.1.pdf - Accessed 19th January 2010.

Khan, A.M. (1981). Urban public transit efficiency: economic and energy factors. Journal of Advanced Transportation, 15(3), 213-230.

Lam, T.N.; Schuler, H.J. (1982). Connectivity index for systemwide transit route and schedule performance. Transportation Research Record, 854, 17-23.

Lee, D. (1989). Transit cost and performance measurement. Transport Reviews, 9(2), 147-170.

Lem, L.L.; Li, J.-L.; Wachs, M. (1994). Comprehensive transit performance indicators. Working Paper UCTC No. 225, Institute of Transportation Studies, School of Public Policy and Social Research, University of California, Los Angeles. Available at http://www.uctc.net/research/papers/225.pdf - Accessed 19th January 2010.

Mackie, P.; Nash, C. (1982). Efficiency and performance indicators: The case of the bus industry. Public Money, 2(3), 41-44.

Massot, M.-H.; Armoogum, J.; Bonnel, P.; Caubel, D. (2006). Potential for car use reduction through a simulation approach: Paris and Lyon case studies. Transport Reviews, 26(1), 25-42.

May, A.D. (2005). PROSPECTS – Procedures for recommending optimal sustainable planning of European city transport systems – Deliverable 15 – Decision Makers' Guidebook. Available at http://www.ivv.tuwien.ac.at/fileadmin/mediapool-verkehrsplanung/Diverse/Forschung/International/PROSPECTS/DMG_English_Version_2005.pdf - Accessed 19th January 2010.

McCrosson, D.F. (1978). Choosing performance indicators for small transit systems. Transportation Engineering, 48(3), 26-30.

Meyer, M.D.; Miller, E.J. (2001). Urban transportation planning – A decision-oriented approach. Mc Graw-Hill, New York, 642pp.

Miller, J.H. (1980). The use of performance-based methodologies for the allocation of transit operating funds. Traffic Quarterly, 34(4), 555-585.

Miller, G.K.; Kirby, R.F. (1984). A structured approach to monitoring and evaluating public transportation services. Transportation Quarterly, 38(1), 23-46.

Morris, J.M.; Dumble, P.L.; Wigan, M.R. (1979). Accessibility indicators for transport planning. Transportation Research A, 13(2), 91-109.

Musso, A.; Vuchic, V.R. (1988). Characteristics of metro networks and methodology for their evaluation. Transportation Research Record, 1162, 22-33.

Nakanishi, Y. (1997). Bus performance indicators – On-time performance and service regularity. Transportation Research Record, 1571, 3-13.

Oum, T.H.; Waters, W.G.; Yu, C. (1999). A survey of productivity and efficiency measurement in rail transport. Journal of Transport Economics and Policy, 33(1), 9-42.

Polzin, S.E.; Pendyala, R.M.; Navari, S. (2002). Development of time-of-day–based transit accessibility analysis tool. Transportation Research Record, 1799, 35-41.

QUATTRO – Quality approach in tendering urban public transport operations (1998). Final report – Synthesis and Recommendations. Available at http://www.transport-research.info/Upload/Documents/200310/quattro.pdf - Accessed 19th January 2010.

REFIT – Refinement and test of sustainable indicators and tools with regard to European transport (2008). Deliverable 7 – Final report of the REFIT project. Available at http://www.transport-research.info/Upload/Documents/200909/20090918_162142_45873_REFIT%20-%20Final%20Report.pdf - Accessed 19th January 2010.

Ryus, P.; Ausman, J.; Teaf, D.; Cooper, M.; Knoblauch, M. (2000). Development of Florida's transit level-of-service indicator. Transportation Research Record, 1731, 123-129.

Seco, A.J.M.; Gonçalves, J.H.G. (2007). The quality of public transport: relative importance of different performance indicators and their potential to explain modal choice. In: Urban Transport XIII: Urban Transport and the Environment in the 21st Century (C.A. Brebbia, ed.), pp. 313-325, WIT Press, Ashurst, Southampton, UK.

Sterman, B.P.; Schofer, J.L. (1976). Factors affecting reliability of urban bus services. Transportation Engineering Journal, 102(TE1), 147-159.

Takyi, I.K. (1993). A multidimensional methodology for evaluating public transportation services. Transportation Research A, 27A(5), 395-405.

Talley, W.K. (1986). A comparison of two methodologies for selecting transit performance indicators. Transportation, 13(3), 201-210.

Talley, W.K. (1988). An economic theory of the public transit firm. Transportation Research B, 22B(1), 45-54.

Talley, W.K.; Anderson, P.P. (1981). Effectiveness and efficiency in transit performance: a theoretical perspective. Transportation Research A, 15A(6), 431-436.

Talley, W.K.; Becker, A.J. (1982). A single measure for evaluating public transit systems. Transportation Quarterly, 36(3), 423-431.

TCRP – Transit Cooperative Research Program (2003a). A guidebook for developing a transit performance-measurement system. TCRP Report 88, Transportation Research Board, Washington, D.C.

TCRP – Transit Cooperative Research Program (2003b). Transit capacity and quality of service manual – 2nd edition – Part 3: Quality of service. TCRP Report 100, Transportation Research Board, Washington, D.C.

Tulkens, H.; Wunsch, P. (1994). Les performances économiques de la STIB au mois le mois. Schweizerische Zeitschrift für Volkswirtschaft und Statistik / Revue Suisse d'Economie Politique et de Statistique, 130(4), 627-646 (in French).

Vaziri, M.; Deacon, J.A. (1984). Peer comparisons in transit performance evaluation. Transportation Research Record, 961, 13-21.
Yeh, C.-H.; Deng, H.; Chang,Y.-H. (2000). Fuzzy multicriteria analysis for performance evaluation of bus companies. European Journal of Operational Research, 126(3), 459-473.
Yin, Y.; Lam, W.H.K.; Miller, M.A. (2004). A simulation-based reliability assessment approach for congested transit network. Journal of Advanced Transportation, 38(1), 27-44.

APPENDIX 1 – LIST OF ABBREVIATIONS USED IN TABLE 2

Acc	Number of accidents
Area	Area where the service is implemented
Cost	Operating costs of the service
Covered area	Area where the service is accessible
Covered pop	Population for which the service is accessible
Deficit	Operating costs minus revenues, excluding subsidies
Empl	Number of employees
Empl h	Employees * hours of work
Exp	Expenses
Fuel	Fuel consumption
Lines length	Total length of public transport lines
Non-lab exp	Expenses excluding labour costs
Op exp	Operating expenses
Op lab exp	Operating labour expenses
Op subs	Subsidies to operating expenses
Paid h	Total hours of work
Pax	Number of passengers
Pax km	Passengers * travelled kilometres
Pax rev	Revenues from passengers
Platform h	Hours of work during service
Pop	Population in the service area
Rev	Operating revenues
Rev veh h	Vehicles * revenue hours of service
Rev veh km	Vehicles * travelled kilometres
Sched veh h	Vehicles * scheduled hours of operations
Seat h	Seats offered * hours of operations
Seat km	Seats offered * travelled kilometres
Subs	Subsidies from public funds
Tot exp	Total expenses
Veh	Number of vehicles
Veh h	Vehicles * hours of operations
Veh km	Vehicles * travelled kilometres
Veh op cost	Operating costs of vehicles